Intermediaries, Inventories and Endogenous Dynamics in Frictional Markets*

Chao Gu
University of Missouri

Lu Wang
University of California-Irvine

Randall Wright
Zhejiang University and Wisconsin School of Business

January 31, 2024

Abstract

We study dynamics in frictional markets with inventories, focusing on models with intermediated trade, where middlemen buy assets or goods from sellers and sell them to buyers. Extending previous work, we include heterogeneous buyer valuations, and develop a characterization of equilibrium in terms of reservation trading strategies (homogeneous valuations imply bang-bang solutions with discontinuities that are awkward for the economics and mathematics). In continuous or discrete time, equilibria exist where market participation, trading strategies, liquidity, and other variables fluctuate as self-fulfilling prophecies. This is driven by strategic considerations, not increasing returns or related assumptions made in other models.

JEL Classification Numbers: E02, E44, G21
Keywords: Inventories, Intermediaries, Dynamics, Search

*For input we thank David Lagakos, Guillame Rocheteau, Michael Choi and Jess Benhabib, as well as participants in presentations at UC-Irvine and the Summer Workshop on Money, Banking, Payments, and Finance at in Gerzensee, Switzerland. Wright acknowledges support from the Ray Zemon Chair in Liquid Assets and the Ken Burdet Chair in Search Theory and Applications at the Wisconsin School of Business. The usual disclaimers apply.

Your inventory cycle is basically a three-step relationship between the supplier who provides the product, the retailer that sells the product, and the consumer who buys the product. Pierre (2019).

1 Introduction

This paper is a study in economic theory, but we think it is has substantial relevance for the way one thinks about markets in the real world. The goal is to understand dynamic equilibria in search-based models of inventories, with a focus on intermediated trade: middlemen buy *assets or goods* from sellers, then sell them to buyers. A key extension over related work is to incorporate heterogeneous valuations and develop a characterization of outcomes in terms of reservation trading strategies, since homogeneous valuations imply bang-bang (corner) solutions that create difficulties for the economics and the mathematics. A main result is the existence of multiple equilibria, including continuous- or discrete-time cycles, where inventories, market participation, liquidity, prices, markups and other variables fluctuate as self-fulfilling prophecies.

Our approach to middlemen and inventories builds on a literature going back to Rubinstein and Wolinsky (1987). This research studies markets with frictions, using search theory, and roles for middlemen arise from their comparative advantage in certain attributes, including matching efficiency, information, bargaining power, and storage cost or capacity.[1] While following in this tradition, past papers mainly concentrate on steady states, or sometimes transitions to steady state. We emphasize the possibility of endogenous fluctuations.

The objects being traded can be either assets or goods, the difference being that inventories of assets yield positive returns, while inventories of goods yield negative returns, i.e., storage costs. That distinction, which is convenient for keeping track of different cases, is borrowed from Nosal et al. (2019), and in that model it makes a big difference, while here it turns out to be less important. Still, interpreting agents

[1] Rather than review the literature, we refer to Wright and Wong (2015), which lists papers where middlemen have advantages in search, information etc. Work since then includes Nosal et al. (2019), Farboodi et al. (2018), Farboodi et al. (2023), Hugonnier et al. (2020) and Gong (2023).

as trading assets is interesting because it connects to the literature on search-based models of OTC (over the counter) financial markets following Duffie et al. (2005). However, although that framework typically has dealers intermediating between asset sellers and buyers, they hold no inventories, and instead simply transfer assets between agents using a frictionless interdealer market (with exceptions, like Weill 2007, but he does not address the issues studied here).

Multiplicity emerges from heterogenous valuations by buyers combined with entry by sellers. When a middleman contacts a buyer, even if there are myopic gains from trade, there is an option to hold out for a future buyer with higher valuation. Agents use reservation strategies, but compared to standard formulations a new strategic effect arises: If middlemen are more inclined to sell to buyers – i.e., have a low reservation value – they more often need to replenish inventories. That makes it easier for sellers to trade, increasing seller entry and making it easier for middlemen to replenish inventories, which rationalizes a low reservation value. However, if reservation values are high, there are few sales to buyers and less need for middlemen to replenish inventories. That reduces seller entry, making it hard to replenish inventories, which rationalizes a high reservation value.

The complementarity between reservation strategies and market participation is novel and is a good reason to have heterogeneous valuations. In a similar model with homogeneous valuations, Nosal et al. (2019) can get multiplicity only in markets for assets, i.e., only if inventories have positive returns. The reason is simple: with homogeneity middlemen never decline trade with a buyer to hold out for one with higher valuation, but might keep rather than sell inventory for its return *if* that is positive. With heterogeneous valuations, we get multiplicity and interesting dynamics for goods or assets, although it is *easier* with assets, as discussed below, consistent with a long tradition of arguing that financial intermediaries are particularly prone to instability or volatility (see Gu et al. 2023 for a discussion and references to those making that claim).

Having said that, we like that the model can generate interesting results in markets for goods, not just assets. One interpretation is that ours is a model of retail trade, and a stylized fact is that the efficiency/productivity of these markets differs dramatically across economies, as discussed by Lagakos (2016). Multiplicity is consistent with the idea that retail markets in some economies may be stuck in a bad equilibrium, where low efficiency/productivity is a self-fulfilling prophecy.[2]

Going beyond steady states, we are interested in the possibility of endogenous fluctuations. To this end bifurcation theory is used to show there are limit cycles. The use of these methods goes back to Benhabib and Nishimura (1979) in growth theory. Applications in search include Diamond and Fudenberg (1989), who get cycles in Diamond (1982a) *if* the matching technology displays increasing returns, and Mortensen (1999), who gets cycles in a version of Pissarides (2000) *if* the production technology displays increasing returns. One might question the empirical relevance of increasing returns, but that aside, it seems fair to say that these results are driven by mechanical technology specifications that play no role here.[3]

There are many papers with multiplicity and endogenous dynamics in monetary economics (see the surveys by Lagos et al. 2017 and Rocheteau and Nosal 2017 for search-based models, and Azariadis 1993 for other approaches). The economic forces behind those results are different, relying on the notion that what you accept in exchange depends on what others accept. One manifestation of the difference is that our results work through heterogenous valuations and endogenous participation, factors that are not needed for the results in monetary theory.[4] A feature of monetary models is that utility is not (perfectly) transferable. Burdett and Wright (1998) show nonmonetary search models with nontransferable utility can also have multiplicity and dynamics while the same environment with transferable utility

[2]Given his expertise, we quote Lagakos, with permission, from correspondence: "That does sound intriguing – I don't remember seeing a paper that says something like that. I have the impression that even countries of similar income levels often have pretty different retail structure and efficiency. That smells like it could be multiple equilibria."

[3]Other search models with dynamics based on increasing returns or related devices include Howitt and McAfee (1988,1992), Boldrin et al. (1993), Kaplan and Menzio (2016) and Sniekers (2018). Fershtman and Fishman (1992), Burdett and Coles (1998) and Albrecht et al. (2013) are examples of search models with somewhat different dynamics.

[4]Kehoe et al. (1993) and Renero (1988) study cycles in the discrete-time model of Kiyotaki and Wright (1989), although Oberfield and Trachter (2012) show the cycles vanish as period length shrinks, as discussed in Section 4.3; this is one reason we use continuous time in much of this paper. See, e.g., Rocheteau and Wright (2013) and references therein for related work.

cannot, and a nice recent application of this to OTC markets is Martel (2023). This is not relevant in our setup, where utility is transferable.[5]

Motivating general interest in inventories, many people think that they are an important component of business cycles, in part because they are volatile and procyclical. This can be understood with a supply-side story: when productivity is high, it is efficient to produce a lot and keep some as inventory to spread good times into the future. This paper instead concerns a demand-side story: holding productivity constant, when inventories are high production slows because middlemen are not buying. This could make inventories countercyclical *if* there were no shocks, but of course there can be many shocks driving cycles. Hence, countercyclicality here describes not the macro data, but what can happen as a self-fulfilling prophecy about inventories, trading strategies and entry decisions.[6]

Section 2 presents a relatively simple specification, without entry, and shows equilibrium is unique. Section 3 adds seller entry and shows how multiplicity and cycles can emerge. Section 4 explores other topics: welfare; entry by middlemen instead of sellers; discrete-time models; and a version where consumers instead of middlemen hold inventories. Section 5 concludes. Proofs are in the Appendix.

2 The Basic Framework

A continuum of infinitely-lived, risk-neutral agents come in three types, labeled B, S and M, for buyers, sellers and middlemen, with differences detailed below. Type i can participate in a continuous-time, bilateral matching market if they pay entry

[5] Transferable utility means there are no payment frictions. One imlication is that, while in what follows we often describe trade between middlemen and others as being decided by the former, in fact the decision is joint: they trade as long as the *total* surplus is positive.

[6] See Blinder (1990) and references therein for early work advocating the importance of inventories in macro; there is too much later work to review here, but see Khan and Thomas (2007) for an example. To give some idea of what macroeconomists find interesting, consider quarterly U.S. data, 1974-2007 (it is common to stop in 2007 to avoid the financial crisis, but the patterns are similar in a longer sample, even if the magnitudes are somewhat different). Inventories over GDP average 0.56. After taking logs and HP filtering, the sd (standard deviation) of inventories over the sd of GDP is 0.86 and their correlation is 0.73, while the sd of sales over the sd of GDP is 0.19 and their correlation is 0.95. This seems consistent with a supply story – when productivity and output are high, it is efficient to consume some and inventory the rest. While that is interesting, ours is a micro model of a market, not designed to capture macro facts.

cost κ_i, but for now $\kappa_i = 0 \ \forall i$ so everyone participates. Indeed, they participate forever, which is not crucial but simplifies some calculations compared to, e.g., Rubinstein and Wolinsky (1987) where M stays forever while B and S exit after one trade, or Nosal et al. (2017) where everyone exits probabilistically after each trade (see also Vanyos and Wang 2007 or Farboodi et al. 2023). When B and S meet, S can produce an indivisible object x, at 0 cost, that gives B match-specific payoff π, with CDF $F(\pi)$ on $[\underline{\pi}, \overline{\pi}]$. Note that π can be utility if B consumes x, or profit if B uses it as an asset for investment or input for production.

In spirit if not detail this is similar to much work following Duffie et al. (2005). There all agents can store x and individual valuations change over time, determining who wants to buy and sell. We are pursuing versions like that in a companion paper, but here, for tractability, it is better to have dedicated sellers S, who are always the originators of x, and dedicated buyers B, who are always the end users.

In any case, S can produce for M, who can store x in inventory, and may or may not sell it to B when they meet. Hence, M can be interpreted as agents with a storage technology – others cannot store x, for now, but see Section 4.4. There is a flow payoff ρ for M with x in inventory, and we say that x is an asset when $\rho > 0$ (it has a return), while x is a good if $\rho < 0$ (it has a storage cost); this usage is not critical, but helps keep track of different cases. Inventory held by M depreciates by disappearing at rate $\delta \geq 0$. As in many papers in search theory, holdings of x by M are constrained to the set $\{0, 1\}$, which is obviously special, but allows one to make salient points in a succinct way.[7]

Let n_i be the measure of $i = B, S, M$ in the market, which is exogenous for now, and n the measure of M with x in inventory, which is endogenous. There is a standard matching technology: you meet someone at Poisson rate α; and each meeting is a random draw from the population. In particular, if N is the measure

[7] In addition to middlemen papers like Rubinstein and Wolinsky (1987), to pick just a few examples, $\{0, 1\}$ restrictions are imposed in the original search-equilibrium model of Diamond (1982a), many of the monetary models cited in fn. 4, banking models like Cavalcanti and Wallace (1999), OTC asset models like Duffie et al. (2005), labor models like Pissarides (2000) and partnership models like Burdett and Coles (1997). We think this puts us in good company.

of all market participants, the arrival rate of B for both M and S is $\alpha n_b/N$, so M has no advantage over S in search. Given $\pi > 0$, when B and S meet they always trade, since it does not affect their continuation values. Also, when S meets M with $x = 0$ they trade unless $\rho < 0$ and $|\rho|$ is big (more on this below). The interesting question is, when M with $x = 1$ meets B, do they trade? As we will see, the answer depends on fundamentals, including π, and on beliefs.

If j gives x to i, the latter pays p_{ij} determined by bargaining with transferable utility. Thus, if Σ_{ij} is the total surplus available when i and j meet, they trade if $\Sigma_{ij} > 0$, and i's surplus is $\theta_{ij}\Sigma_{ij}$, where $\theta_{ij} \geq 0$ is i's bargaining power against j, with $\theta_{ij} + \theta_{ji} = 1$. Letting $V_{s,t}$ and $V_{b,t}$ be value functions for S and B, $V_{x,t}$ the value function for M with $x \in \{0, 1\}$, and $\Delta_t = V_{1,t} - V_{0,t}$, we have[8]

$$\Sigma_{bs,t} = \pi,\ \Sigma_{ms,t} = \Delta_t, \text{ and } \Sigma_{bm,t} = \pi - \Delta_t. \tag{1}$$

Note the continuation values and threat points for S and B cancel in the surpluses, so V_s and V_b do not appear. From these follow what we call the direct price, the wholesale price, and the retail price, given respectively by

$$p_{bs,t} = \theta_{sb}\pi,\ p_{ms,t} = \theta_{sm}\Delta_t, \text{ and } p_{bm,t} = \theta_{mb}\pi + \theta_{bm}\Delta_t. \tag{2}$$

When M with $x = 1$ meets B with valuation π, they trade with probability $\tau_t = \tau(\pi, R_t)$, where R_t is the reservation value:

$$\tau(\pi, R_t) = \begin{cases} 0 & \text{if } \pi < R_t \\ [0,1] & \text{if } \pi = R_t \\ 1 & \text{if } \pi > R_t \end{cases} \tag{3}$$

Clearly, $R_t = \Delta_t$. Hence, the expected flow payoff for B is

$$rV_{b,t} = \frac{\alpha n_{s,t}}{N_t}\theta_{bs}\mathbb{E}\pi + \frac{\alpha n_t}{N_t}\theta_{bm}\mathbb{E}\left[\tau(\pi, \Delta_t)(\pi - \Delta_t)\right] + \dot{V}_{b,t}, \tag{4}$$

where r is the discount rate and and prices have been eliminated using (2). The first term on the RHS is the arrival rate of S times B's share of the surplus; the

[8] At this point we start subscripting variables by t, including n_s, N and n_m even though they are constant in this most basic version of the environment, so that the same expressions hold when they are endogenous; we do not subscript n_b by t since it is fixed in all versions.

second is the arrival rate of M with $x = 1$ times the probability they trade times B's share of the surplus; the third is the pure time change in value.

Similarly, for S,

$$rV_{s,t} = \frac{\alpha n_b}{N_t}\theta_{sb}\mathbb{E}\pi + \frac{\alpha\left(n_{m,t} - n_t\right)}{N_t}\theta_{sm}\Delta_t + \dot{V}_{s,t}, \tag{5}$$

and for M,

$$rV_{0,t} = \frac{\alpha n_{s,t}}{N_t}\theta_{ms}\Delta_t + \dot{V}_{0,t} \tag{6}$$

$$rV_{1,t} = \frac{\alpha n_b}{N_t}\theta_{mb}\int_{\Delta_t}^{\infty}\left(\pi - \Delta_t\right)dF\left(\pi\right) + \rho - \delta\Delta_t + \dot{V}_{1,t}. \tag{7}$$

Subtracting (7)-(6) and simplifying with integration by parts, we get

$$\dot{\Delta}_t = -\frac{\alpha n_b}{N_t}\theta_{mb}\int_{\Delta_t}^{\infty}\left[1 - F\left(\pi\right)\right]d\pi + \frac{\alpha n_{s,t}}{N_t}\theta_{ms}\Delta_t - \rho + \left(r + \delta\right)\Delta_t. \tag{8}$$

The evolution of inventories held by M is

$$\dot{n}_t = \frac{\alpha n_{s,t}\left(n_{m,t} - n_t\right)}{N_t} - \frac{\alpha n_b n_t \mathbb{E}\tau\left(\pi, \Delta_t\right)}{N_t} - \delta n_t, \tag{9}$$

where $\mathbb{E}\tau\left(\pi, \Delta\right) = \Pr\left(\pi > \Delta\right)$ is the unconditional probability that M and B trade. The first term on the RHS is the measure of M without x times the rate at which they buy it from S; the second is the measure of M with x times the rate at which they sell it to B; the third is depreciation.

Equilibrium is defined as a path for (Δ_t, n_t) satisfying dynamical system (8)-(9), plus the standard side condition that the paths must be nonnegative and bounded, and the initial condition n_0 giving inventories at $t = 0$. A steady state is a constant (Δ, n) satisfying (8)-(9). Given an equilibrium (Δ_t, n_t), or a steady state (Δ, n), all other variables follow, including payoffs, prices, trade volume, etc.

With no intermediaries, $n_m = 0$, equilibrium is obviously unique, with B and S trading whenever they meet. With $n_m > 0$, first notice that the path of Δ_t is independent of n_t. Then from (8)

$$\frac{d\dot{\Delta}_t}{d\Delta_t} = \frac{\alpha n_b}{N_t}\theta_{mb}\left[1 - F\left(\Delta_t\right)\right] + \frac{\alpha n_{s,t}}{N_t}\theta_{ms} + r + \delta > 0,$$

implying Δ_t must equal its steady state value $\forall t$, since any other solution to (8) diverges – a result that reappears in some, but not all, formulations below, and is discussed more later. Given Δ, (9) implies

$$\frac{d\dot{n}_t}{dn_t} = -\left[\frac{\alpha n_{s,t}}{N_t} + \frac{n_b \mathbb{E}\tau(\pi,\Delta)}{N_t} + \delta\right] < 0,$$

so n_t converges monotonically to its steady state. Equilibrium is unique. There are no dynamics due to self-fulfilling expectations.

3 The Main Model

Now let S face a participation decision, which is natural, and nice because it lets us compare economies with and without middlemen while keeping the environment otherwise the same.[9] Then $n_{s,t}$ and N_t can vary with time, while n_m and n_b are constant. The entry condition $rV_{s,t} = \kappa_s$ implies $\dot{V}_{s,t} = 0$. Combine the entry condition with (5) to get

$$N_t \kappa_s = \alpha n_b \theta_{sb} \mathbb{E}\pi + \alpha (n_m - n_t) \theta_{sm} \Delta_t. \tag{10}$$

This lets us eliminate N_t from (8) and (9), resulting in a two-dimensional system

$$\begin{bmatrix} \dot{\Delta}_t \\ \dot{n}_t \end{bmatrix} = \begin{bmatrix} f(n_t,\Delta_t) \\ g(n_t,\Delta_t) \end{bmatrix}, \tag{11}$$

where

$$\begin{aligned} f(n_t,\Delta_t) &= -\frac{\alpha n_b \kappa_s}{\alpha n_b \theta_{sb}\mathbb{E}\pi + \alpha (n_m - n_t)\theta_{sm}\Delta_t}\theta_{mb}\int_{\Delta_t}^{\infty}[1-F(\pi)]\,d\pi \\ &\quad +\alpha\left[1-\frac{(n_m+n_b)\kappa_s}{\alpha n_b \theta_{sb}\mathbb{E}\pi + \alpha (n_m - n_t)\theta_{sm}\Delta_t}\right]\theta_{ms}\Delta_t - \rho + (r+\delta)\Delta_t. \\ g(n_t,\Delta_t) &= \alpha\left[1-\frac{(n_m+n_b)\kappa_s}{\alpha n_b \theta_{sb}\mathbb{E}\pi + \alpha (n_m - n_t)\theta_{sm}\Delta_t}\right](n_m - n_t) \\ &\quad -\frac{\alpha n_b n_t \mathbb{E}\tau(\pi,\Delta_t)\kappa_s}{\alpha n_b \theta_{sb}\mathbb{E}\pi + \alpha (n_m - n_t)\theta_{sm}\Delta_t} - \delta n_t \end{aligned}$$

[9] The environment is the same with and without M in the sense that it always has endogenous market composition due to S entry. With M entry, we eliminate endogenous composition if we eliminate M, but that case is still covered in Section 4.2; we tried entry by B, too, but it is less interesting, unsurprisingly, since type B is fairly mechanical here.

Define the n locus and Δ locus as the curves in (n, Δ) space along which $\dot{n} = 0$ and $\dot{\Delta} = 0$, with their intersections constituting steady states. Their slopes are given by

$$\left.\frac{d\Delta}{dn}\right|_{\dot{\Delta}=0} = \frac{\theta_{sm}\Delta\left\{n_b\theta_{mb}\int_{\Delta}^{\infty}\left[1-F(\pi)\right]d\pi+(n_m+n_b)\,\theta_{ms}\Delta\right\}}{N\kappa_s D} \tag{12}$$

$$\left.\frac{d\Delta}{dn}\right|_{\dot{n}=0} = \frac{n_s+\frac{1}{\kappa_s}\theta_{sm}\Delta\left[\alpha\left(n_m-n\right)-\delta n\right]+n_b\left[1-F(\Delta)\right]+N\frac{\delta}{\alpha}}{\frac{n_m-n}{\kappa_s}\theta_{sm}\left[\alpha\left(n_m-n\right)-\delta n\right]+n_b n f(\Delta)} \tag{13}$$

where

$$\begin{aligned} D &= \frac{(n_m-n)\,\theta_{sm}}{N\kappa_s}\left\{n_b\theta_{mb}\int_{\Delta}^{\infty}\left[1-F(\pi)\right]d\pi+(n_m+n_b)\,\theta_{ms}\Delta\right\} \\ &\quad +n_b\theta_{mb}\left[1-F(\Delta)\right]/\alpha+N(r+\delta)/\alpha^2+n_s\theta_{ms}/\alpha, \end{aligned}$$

As both slopes are positive, there are potentially multiple steady states. This can be illustrated in the degenerate case where $\pi = \bar{\pi}$ with probability 1, even if we are less interested in that than heterogenous valuations. In this the degenerate case, there are three possible *regimes*: (i) $\Delta < \bar{\pi}$, so M with $x = 1$ and B trade with probability $\tau = 1$; (ii) $\Delta > \bar{\pi}$, so they trade with probability $\tau = 0$; and (iii) $\Delta = \bar{\pi}$, and they trade with probability $\tau \in (0, 1)$.

Consider first $\rho > 0$, which can be interpreted as x being an asset rather than a good. We have the following result (all proofs are in the Appendix):

Proposition 1 *Suppose $\pi = \bar{\pi}$ with probability 1 and $\rho > 0$. There exists $\tilde{\rho} > 0$ and $\hat{\rho} > \tilde{\rho}$ such that: (i) if $\rho \in [0, \tilde{\rho})$ there is a unique steady state and it has $\Delta < \bar{\pi}$; (ii) if $\rho \in (\hat{\rho}, \infty)$ there is a unique steady state and it has $\Delta > \bar{\pi}$; (iii) if $\rho \in (\tilde{\rho}, \hat{\rho})$ there are three steady states, $\Delta < \bar{\pi}$, $\Delta > \bar{\pi}$, and $\Delta = \bar{\pi}$.*

Example 1: $\alpha = 1$, $\delta = 0.008$, $r = 0.04$, $n_b = 0.05$, $n_m = 0.5$, $\theta_{mb} = 0.7$, $\theta_{sb} = 1$, $\theta_{sm} = 0.5$, $\kappa_s = 0.1$, $\bar{\pi} = 1$, and various ρ.

Fig. 1 illustrates the result for Example 1. As the left panel shows, for $\rho = 0.1$ there is one steady state; for $\rho = 0.2$ there are three; and for $\rho = 0.3$, there is one. The right panel is for a discrete-time version of the same specification, as analyzed in Section 4.3; it can be ignored for now, but it is perhaps interesting to see the continuous and discrete formulation side by side.

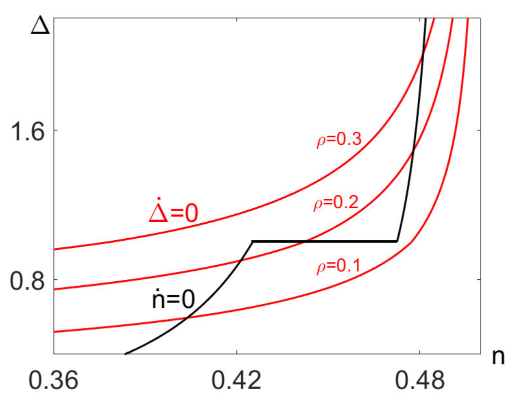

Fig. 1a: Example 1.

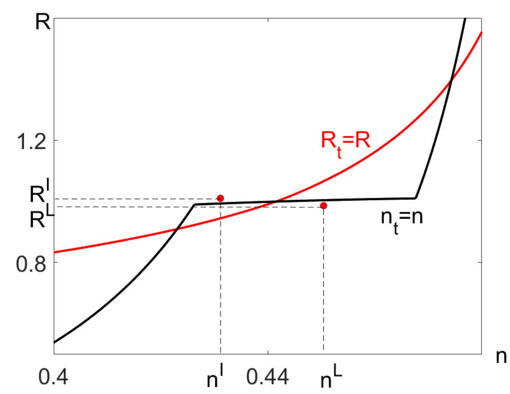

Fig. 1b: Discrete-time version.

Multiple steady states can be explained as discussed in the Introduction: If Δ is low, M with x trades it to B, so the probability M has x is low, which encourages S entry, making n_s high and making it easy for M to get x, consistent with low Δ. If Δ is high, M with x does not trade it to B, so the probability M has x is high, which discourages S entry, making n_s low, consistent with high Δ. When both $\tau = 1$ and $\tau = 0$ are consistent with equilibrium, as usual, so is some $\tau \in (0, 1)$.

Market liquidity – the ease with which agents can buy and sell x – is high (low) when Δ is low (high). Multiplicity means market liquidity is not pinned down by fundamentals. Notice $\tilde{\rho} > 0$ in Proposition 1, which means steady state is unique for $\rho = 0$. This is also true for $\rho < 0$, which says that goods markets have a unique steady state when valuations are homogeneous:[10]

Proposition 2 *Suppose* $\pi = \bar{\pi}$ *with probability* 1 *and* $\rho < 0$. *Then there is a unique steady state and it has* $\Delta < \bar{\pi}$.

[10] The system in the text is derived assuming M and S trade. That is fine if $\rho > 0$, or if $\rho < 0$ and $|\rho|$ is not too big, since then steady state is consistent with $\Delta > 0$, so M and S must trade. However, if $\rho < 0$ and $|\rho|$ is too big, M and S will not trade, and any M with x would dispose of it; in this case, in the only steady state M does not trade. For now, if $\rho < 0$ we simply assume $|\rho|$ is not too big, and return to the issue in Section 4.3.

While our specification is somewhat different, the above results are consistent with Nosal et al. (2019), but things change when we depart from a degenerate π. First, if π is degenerate, multiplicity cannot arise if $\rho \leq 0$ because M's only alternative to trading with B is to keep x, but this "buy and hold" strategy only makes sense if $\rho > 0$. With nondegenerate π there is a different motive for M to pass on trade with B: if the match-specific π is low, M may want to hold out for a higher π. This is standard fare in search theory, and does not rely on $\rho > 0$, although higher ρ helps in the same way that, say, higher unemployment benefits help support higher reservation wages in labor markets. So multiplicity does not require $\rho > 0$ but one can say that $\rho > 0$ might make it more likely.

There are other reasons to go beyond degenerate π. First, when there are multiple steady states, as in Fig. 1, it is hard to characterize dynamics around the middle one because the n locus is horizontal.[11] Also, with disperse π, M and B are only indifferent to trade in the rare event $\pi = \Delta$ (i.e., most of their interactions are strict best responses). Also, if Δ varies over time, intermediation activity varies, but with degenerate π we get bang-bang situations (i.e., τ is almost always 0 or 1), while with disperse π we can get trade volume varying smoothly over time. All of this is verified below, starting with the claim that multiplicity can arise with $\rho < 0$ once π is nondegenerate.

Example 2: $\alpha = 0.96$, $\delta = 0.001$, $r = 0.01$, $n_b = 0.055$, $n_m = 0.4$, $\theta_{mb} = 0.95$, $\theta_{sb} = 1$, $\theta_{sm} = 0.1$, $\kappa_s = 0.225$, $\rho = -0.001$ and

$$F(\pi) = \begin{cases} y_1 \pi / a & \text{if } 0 \leq \pi \leq a \\ y_1 + (y_2 - y_1)(\pi - a)/(b-a) & \text{if } a < \pi \leq b \\ y_2 + (y_3 - y_2)(\pi - b)/(c-b) & \text{if } b < \pi \leq c \\ y_3 + (1 - y_3)(\pi - c)/(d-c) & \text{if } c < \pi \leq d \end{cases} \tag{14}$$

with $a = 1$, $b = 1.02$, $c = 2.38$, $d = 2.4$, $y_1 = 0.01$, $y_2 = 0.5$, and $y_3 = 0.55$.

Fig. 2a shows the situation for Example 2, with $\rho < 0$ and a distribution of π that is disperse but concentrated around two values, as is useful for making a point

[11] As Nosal et al. (2019) say: "Whether this [equilibrium] converges to steady state, or to a small cycle around it, is hard to say from the numerical output, and checking local stability directly is hindered by the [relevant objects] being nondifferentiable."

even if the examples below use a simpler (uniform) distribution of π. The point is that there are three steady states, $(0.156, 0.966)$, $(0.177, 1.012)$ and $(0.198, 1.071)$. Hence, multiplicity obtains with $\rho < 0$, with intuition similar to that laid out above: if Δ is high M only trades with B when π is high, so the probability M has x is high, which reduces entry by S and makes it hard to M to replenish inventory, justifying a high Δ; while if Δ is low, and so on.

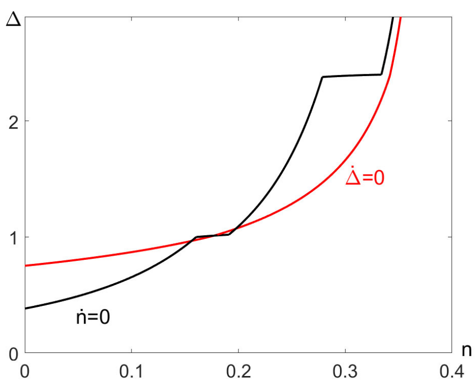

Fig. 2a: Example 2, Steady States.

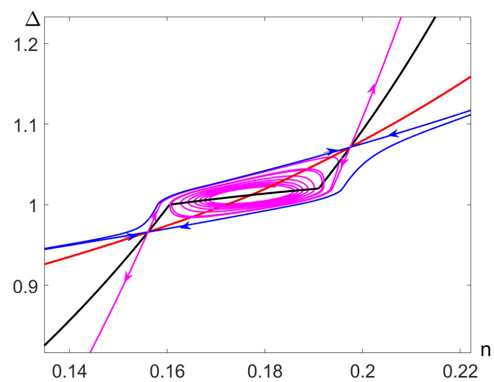

Fig. 2b: Example 2, Phase Plane.

As mentioned, higher Δ means the market is less liquid. To consider related variables, the (average) markup is the ratio of the retail and wholesale prices,

$$\sigma = \frac{\frac{\int_\Delta^\infty p_{bm} dF(\pi)}{1-F(\Delta)}}{p_{ms}} = \frac{\int_\Delta^\infty (\theta_{mb}\pi + \theta_{bm}\Delta)\, dF(\pi)}{\theta_{sm}\Delta\,[1-F(\Delta)]}.$$

The spread is the difference between these prices,

$$\mu = \frac{\int_\Delta^\infty (\theta_{mb}\pi + \theta_{bm}\Delta)\, dF(\pi)}{1-F(\Delta)} - \theta_{sm}\Delta.$$

Trade volume is

$$\upsilon = \frac{\alpha n_b n_s}{N} + \frac{\alpha n_b n}{N}\int_\Delta^\infty dF(\pi) + \frac{\alpha n_s (n_m - n)}{N}.$$

These are relevant because the markup, spread or volume are used as measures of frictions in decentralized markets in both theory (e.g., Weill 2008; Lagos and Rocheteau 2009) and empirical work (e.g., Brennan et al. 1998).

Across steady states in Fig. 2a, the markup, spread and volume are $(\sigma, \mu, \upsilon) = (16.941, 1.540, 0.0372)$, $(18.835, 1.805, 0.030)$ and $(21.138, 2.157, 0.025)$. At higher Δ, both retail and wholesale prices are higher, but on net the markup and spread rise, while volume falls, as might be expected in a less liquid market. Later we check how these variables behave over time, not just across steady state.

To begin the dynamic analysis, consider Fig. 2b, which zooms in around the three steady states. It is easy to check that the lower and upper steady states are saddle points, and their stable (unstable) manifolds are shown in blue (pink). For these parameter values the middle steady state is a sink, with branches of the unstable manifolds of the other steady states spiraling in towards it. There are multiple dynamic equilibria: starting from any n_0 in some range, equilibrium can converge to the upper or lower steady state, or it can spiral into the middle steady state, and what happens depends on initial beliefs as given by Δ_0.

Example 3 (saddle loop bifurcation): $\pi \sim U[0, 2]$, $\alpha = 1$, $\delta = 0.001$, $n_b = 0.05$, $n_m = 0.5$, $\theta_{mb} = 0.75$, $\theta_{sb} = 1$, $\theta_{sm} = 0.05$, $\kappa_s = 0.1$, $\rho = 0.108$ and various r.

Fig. 3a shows the situation for Example 3, where again there are three steady states, but now the middle one is a source, not a sink. Again, starting from any n_0 in some range, there are many equilibria depending on Δ_0, but now we cannot spiral into the middle steady state. This suggests the possibility of cycles which we now explore using bifurcation theory.[12]

The first case involves a saddle loop (also called a homoclinic) bifurcation. In Fig. 3a, with $r = 0.018$, the blue stable manifold going to the lower steady state is inside the pink unstable manifold. In Fig. 3b, with $r = 0.013$, the blue stable manifold is outside the pink unstable manifold. By continuity, for some $r^* \in (0.013, 0.018)$ there exists a homoclinic orbit – i.e., the unstable and stable manifold

[12] References on the dynamical system theory used here include Guckenheimer and Holmes (1983) and Kuznetsov (2004), while Azariadis (1993) is a standard source for economic applications. We employ the Hopf bifurcation, as used to get continuous-time cycles in a search model by, e.g., Diamond and Fudenberg (1989), and the saddle loop bifurcation, used by, e.g., Coles and Wright (1998) or Mortensen (1999). Sniekers (2018) uses the Bogdanov-Takens bifurcation, not previously used in search theory, but used in a macro model by Benhabib et al. (2001). While Sniekers (2018) approach may have some advantages, we find it less tracatbale, and in any case we get what we are after with our approach.

coincide – as shown in Fig. 3c. As the middle steady state inside the homoclinic orbit is a source for these parameters, and other obits inside it cannot get out, the inescapable conclusion is this: starting inside the homoclinic orbit, since the system cannot escape and cannot go to the middle steady state, it must go to a cycle. The green curve in Figure 3d, drawn for $r = 0.016$, is a trajectory starting near the middle steady state, and the pink curve is the unstable manifold of the lower steady state, both of which approach a limit cycle.

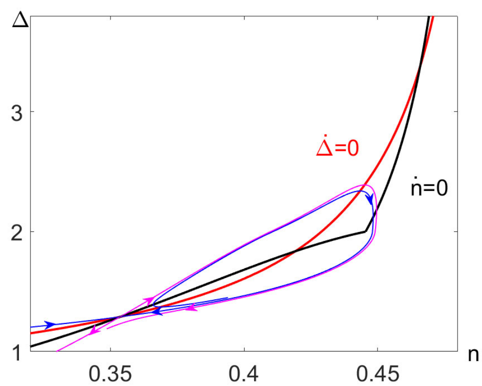

Fig. 3a: Example 3, $r = 0.018$.

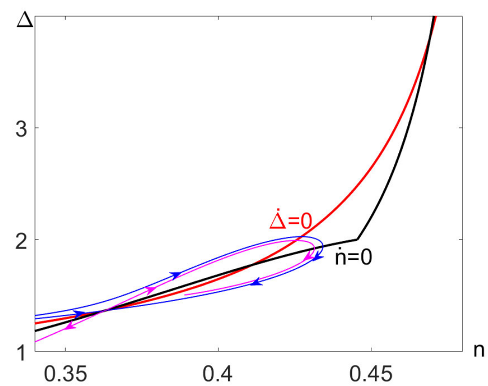

Fig. 3b: Example 3, $r = 0.013$

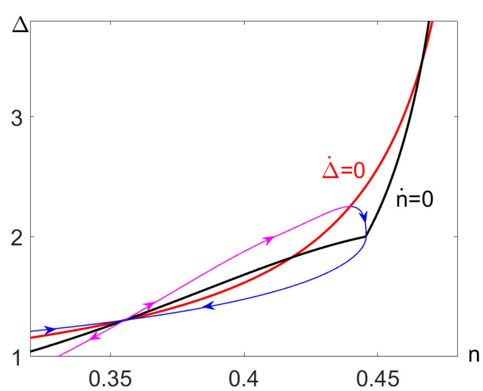

Fig. 3c: Example 3, homoclinic orbit

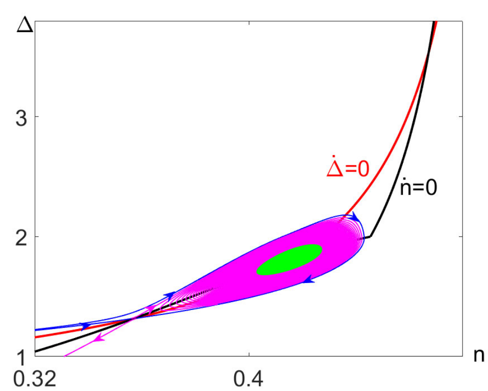

Fig. 3d: Example 3, $r = 0.016$.

The mechanics of saddle loop bifurcations are clear from the graphs, but more formally the Andronov-Leontovich theorem says: Consider a system $\dot{\mathbf{x}} = f(\mathbf{x}, r)$ with $\mathbf{x} \in \mathbb{R}^2$ and a parameter $r \in \mathbb{R}^1$ where f is smooth. Suppose at $r = r^*$ there is a steady state x^* that is a saddle point that has a homoclinic orbit with another steady state inside it. Under mild regularity conditions (Kuznetsov 2004, Section 6.2), $\forall r$ in a nondegenerate neighborhood of r^* there exists a neighborhood of the homoclinic orbit and x^* in which a unique limit cycle bifurcates from the homoclinic orbit (i.e., the cycle emerges as r crosses r^*). The theorem also gives conditions under which cycles are stable or unstable, but the result to emphasize is that they exist for all r in a neighborhood of r^* – i.e., for a range of parameter with positive measure – even if the homoclinic orbit exists only at r^*.

Time series from this cycle are shown in Fig. 3e. While the examples are not meant to be calibrations, only to show mathematical possibilities, we mention that with $r = 0.016$ a period corresponds to roughly 1 quarter, giving the cycle a not-unrealistic duration of about 7 years. Having said that, there is as usual a tradeoff in the sense that higher frequency cycles generally have lower amplitude. In any case, notice entry and volume lead Δ, while inventories and output lag, Δ. Also, the markup (spread) is negatively (positively) correlated with Δ.

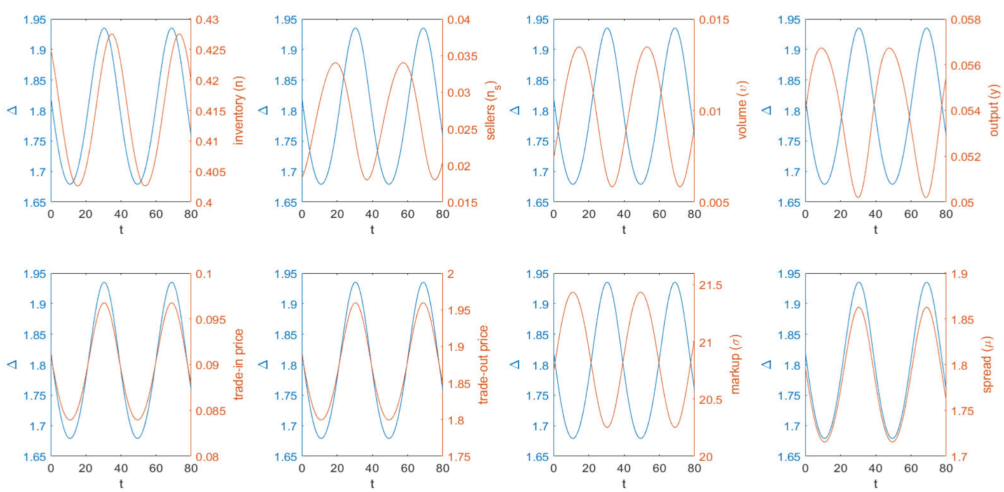

Fig. 3e: Example 3, $r = 0.016$, Time Series.

Example 4 (Hopf bifurcation, subcritical): $\pi \sim U[0, 3]$, $\alpha = 1$, $\delta = 0.0001$, $r = 0.0825$, $n_b = n_m = 1$, $\theta_{mb} = \theta_{sb} = \theta_{sm} = 0.5$, $\kappa_s = 0.4$ and $\rho = 0.33$.

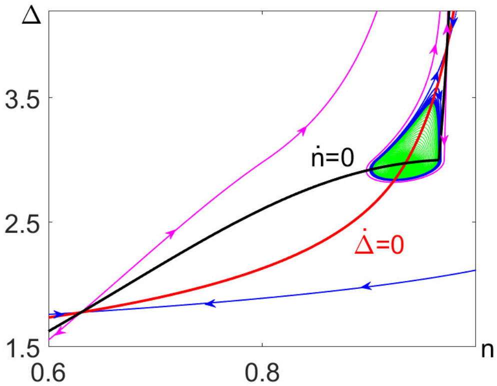

Fig. 4a: Example 4, $r = 0.0825$.

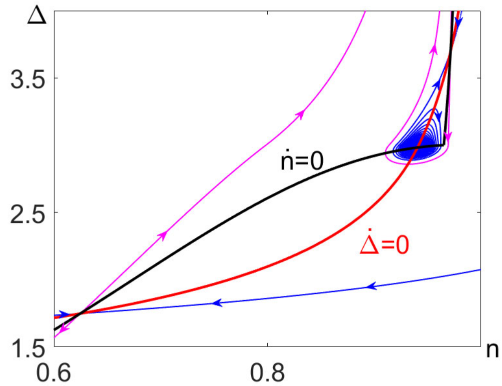

Fig. 4b: Example 4, $r = 0.0875$.

An alternative approach uses the Hopf bifurcation. Fig. 4 is for Example 4, again with three steady states, and blue (pink) curves showing the stable (unstable) manifolds. The middle steady state can be a sink or a source. As r increases there is a Hopf bifurcation at $r^* = 0.0851$ where the trace of the system is 0: for $r < r^*$ the middle steady state is a sink; for $r > r^*$ it is a source. With $r = 0.0825$ in Fig. 4a, the stable manifold spirals away from an unstable cycle and goes to the upper steady state, and shown in green is a trajectory spiraling away from the cycle toward the sink. As r increases above r^* the sink becomes a source and the cycle disappears, as shown in Fig. 4b for $r = 0.0875$. In this example the bifurcation is subcritical, meaning a small increase in r around r^* can cause the system to deviate away from the middle steady state.

Fig. 4c plots time series with $r = 0.0825$. Volume, output and entry by S are negatively correlated with Δ, while inventories are positively correlated with Δ. Notice that over the cycle M and B trade with positive probability when $\Delta < 3$, where $\bar{\pi} = 3$ is the upper bound of the support, and do not trade at all when $\Delta > 3$. This can be described as recurrent intermediation freezes and thaws.[13] The market does not shut down during a freeze, since B still trade with S, but B cannot trade

[13] See Gu et al. (2022) and references therein for a discussion of freezes in asset and credit markets, as well as attempts to model them formally that are very different from our approach – e.g., using stochastic (sunspot) equilibrium in discrete time.

with M, who are rationally holding out for better times. Since M does not trade with B during a freeze, Fig. 4c only shows the markup and spread during thaws.

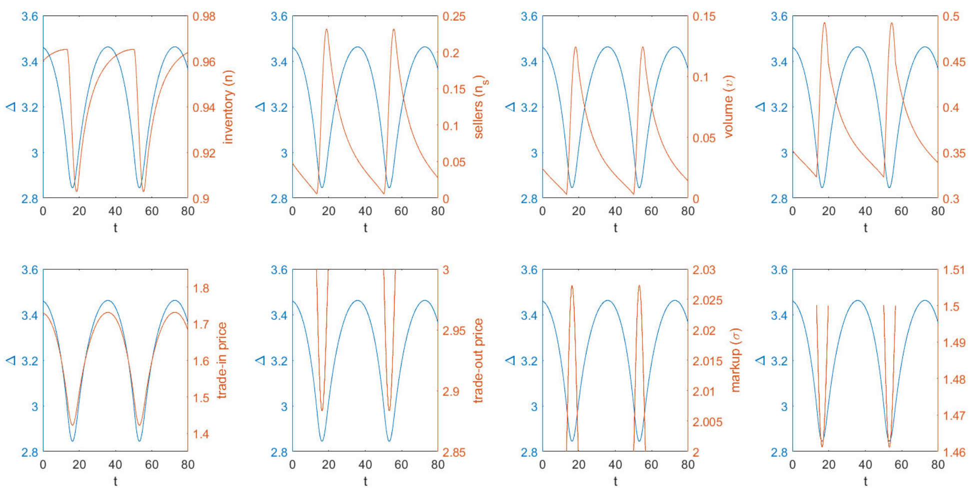

Fig. 4c: Example 4 Time Series.

Example 5 (Hopf bifurcation, supercritical): $\pi \sim U[0,2]$, $\alpha = 1$, $\delta = 10^{-5}$, $n_b = 2$, $n_m = 1$, $\theta_{mb} = 0$, $\theta_{sb} = 1$, $\theta_{sm} = 0.2$, $\kappa_s = 0.6$ and $\rho = 0.3$.

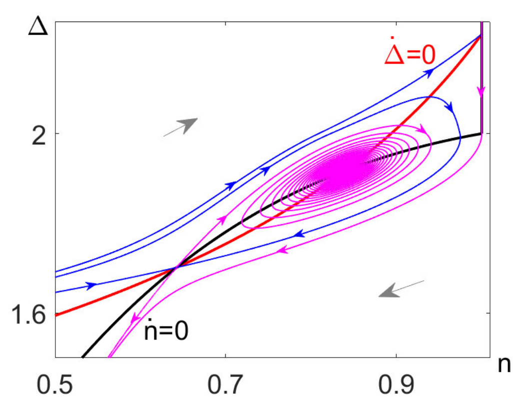

Fig 5a: Example 5, $r = 0.055$.

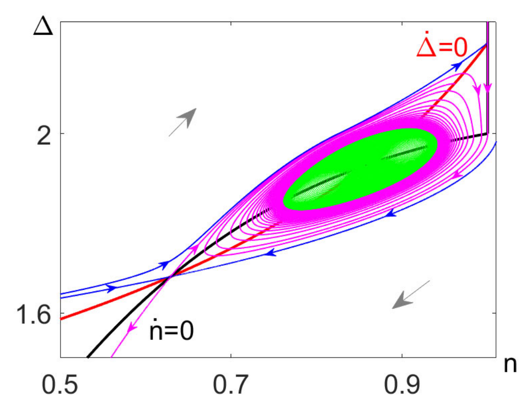

Fig. 5b: Example 5, $r = 0.0562$.

Fig 5 is for Example 5, with a bifurcation at $r^* = 0.0557$. In Fig 5a, with $r = 0.055$, the middle steady state is a sink and the unstable manifold of the lower steady state converges to it. As r rises past r^* the sink becomes a source with a stable limit cycle around it. In Fig. 5b, with $r = 0.0562$, the green curve is a trajectory spiraling away from the source, converging to a cycle. The unstable manifolds also converge to a cycle. Fig. 5c plots time series, like Fig. 4c, with a few differences – e.g., the variability of the markup is smaller, and while there are again freezes, they are shorter, and the series are smoother. Also, similar to the saddle loop, with a Hopf bifurcation cycles exist for a set of parameters with positive measure, not just at the bifurcation point r^* (Kutznesov 2004, Theorem 3.4).

In terms of economics, these examples show it is not hard to find parametric specifications with cyclic equilibria due to self-fulfilling expectations about trading and entry decisions. We do not claim that actual data are best explained by such cycles in isolation – presumably observations from the real world are driven at least in part by fundamentals, including shocks to technology, policy, etc. We do suggest this: when simple models deliver equilibria with endogenous variables fluctuating purely due to beliefs, it lends credence to the idea that markets in real economies might be susceptible to similar phenomena. Therefore we think it is useful to analyze models with natural ingredients, like inventories and entry decisions, to see if and when they display such phenomena.

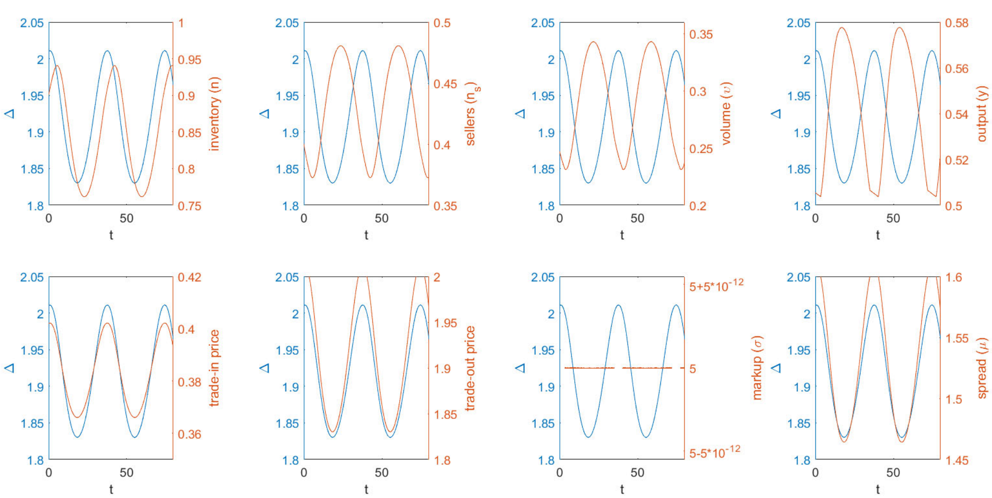

Fig. 5c: Example 5, $r = 0.0562$, Time Series.

4 Other Issues

4.1 Welfare

In the above specification, with entry by S, steady state welfare is:

$$\begin{aligned} W &= r\left[n_b V_b + (n_m - n) V_1 + n V_0\right] \\ &= \frac{\alpha n_s n_b}{N}\theta_{bs}\mathbb{E}\pi + \frac{\alpha n n_b}{N}\int_{\Delta}^{\infty}(\pi - \Delta)\, dF(\pi) + \frac{\alpha n_s (n_m - n)}{N}\theta_{ms}\Delta + n(\rho - \delta\Delta) \end{aligned}$$

Thus, W includes the surplus when B trades with S, when B trades with M and when M trades with S, plus the flow payoff from dividends minus the loss due to depreciation.

Different steady states are distinguished by their Δ, with higher Δ reducing entry. The first term falls with Δ because the number of meetings between B and S falls. The second term is ambiguous because while the surplus in these meeting falls the number of meetings can go either way. The third term is also ambiguous because the number of meetings can go either way. The last term is ambiguous because the total dividend and the total depreciated value both increase in Δ. In most of our examples, W decreases with Δ, but in Example 5 with $r = 0.0562$, W increases in Δ.

Whether welfare is lower on the cycle depends on parameter values as well as where the cycle starts. Consider a case with three steady states, let W_L, W_M and W_U be welfare in the lower, middle and upper ones, and let W_C be welfare in a cycle. In Example 3, if a cycle starts at the highest Δ then $W_L > W_M > W_C > W_U$, while if it starts at the lowest Δ then $W_L > W_C > W_M > W_U$. In Example 5, if a cycle starts at the highest Δ then $W_U > W_C > W_M > W_L$, while if it starts at the lowest Δ then $W_U > W_M > W_C > W_L$. The conclusion is that welfare comparisons are generally ambiguous, as is no surprise, based on previous work in the area.

In terms of efficiency of intermediation, in general, W can be higher or lower with middlemen than without middlemen, as in, e.g., Nosal et al. (2015,2019), Masters (2007,2008), Farboodi et al. (2019) and Gong (2023). The reason is that while M perform a real service, their activity depends on bargaining power, and they like to buy low and sell high; hence, they may operate even if it is not socially efficient.[14]

4.2 Entry by Middlemen

With endogenous participation by M instead of S, the dynamical system is described by (4)-(9) with constant n_s, time-varying $n_{m,t}$ and entry condition $rV_{0,t} = (n_s/N_t)\,\alpha\theta_{ms}\Delta_t = \kappa_m$. Combining this condition with (8) we get

$$\dot{\Delta} = -\frac{\kappa_m \theta_{mb} n_b}{\theta_{ms} n_s \Delta} \int_{\Delta}^{\infty} [1 - F(\pi)]\, d\pi + \kappa_m - \rho + (r + \delta)\,\Delta.$$

This is a first-order differential equation, with $d\dot{\Delta}/d\Delta > 0$. Hence, the results are similar to the version with no entry in Section 2: the unique equilibrium has Δ constant at its steady state level. Also,

$$\dot{n} = \alpha n_s - \frac{\kappa_m (n_b + n_s + n)}{\theta_{ms}\Delta} - \frac{\kappa_m n_b n\,[1 - F(\Delta)]}{n_s \theta_{ms} \Delta} - \delta n_t,$$

which implies $d\dot{n}/dn < 0$, so n_t converges to steady state.

While n_t adjusts during the transition Δ_t does not change, the way payoffs do not vary in Pissarides (2000) even while unemployment adjusts to steady state.[15] We can also relate the results to Rocheteau and Wright (2005), where buyers choose money balances before entering the market. If seller entry is endogenous, there can be multiple equilibria, since there is a complementarity between buyer and seller strategies; but if buyer rather than seller entry is endogenous, there cannot be multiple equilibria since, heuristically, the same agents are making the money holding and entry decisions. A similar intuition applies here, although the mechanism is different, since it is the complementary between the trading strategy of M and entry decision of S that matters. Still, the idea is that there is no multiplicity with

[14] We do not pursue welfare further because it has been studied elswhere, and because the results are similarly ambiguous. Still, to illustrate the possibilities, let W_0 be welfare with $n_m = 0$, including only the surplus from trade between B and S, and consider a case with three steady states. Letting W_L, W_M and W_U be as in the text, we have: In Example 1, $W_L > W_M > W_U > W_0$. In Example 4 with $r = 0.0825$, $W_L > W_M > W_0 > W_U$. There are also examples with $W_L > W_0 > W_M > W_U$. In Example 5 with $r = 0.0562$, $W_U > W_M > W_L > W_0$, which shows that, when ρ is high, more inventories and a less liquid market can entail higher welfare.

[15] While this is also true in Section 2, the economics is perhaps more clear here because entry by M is similar to entry (vacancy posting) by firms in Pissarides (2000).

entry by M since the same agents are making the trading and entry decisions (more on this below).

4.3 Discrete Time

Now consider a discrete-time model, with α the meeting probability, δ the depreciation probability and $\beta \in (0,1)$ the discount factor. The surpluses are

$$\Sigma_{bs,t} = \pi,\ \Sigma_{ms,t} = (1-\delta)\,\beta\Delta_{t+1},\ \Sigma_{bm,t} = \pi - (1-\delta)\,\beta\Delta_{t+1},$$

where again $\Delta_t = V_{1,t} - V_{0,t}$. Now $R_t \equiv (1-\delta)\,\beta\Delta_{t+1}$ is the reservation value satisfying $\Sigma_{mb,t} = 0$. Prices are as in (2) except R_t replaces Δ_t.

Letting $\tau(\pi, R_t)$ be as in (3), the discrete-time value functions are

$$V_{b,t} = \frac{\alpha n_{s,t}}{N_t}\theta_{bs}\mathbb{E}\pi + \frac{\alpha n_t}{N_t}\tau(\pi, R_t)\,\theta_{bm}(\pi - R_t) + \beta V_{b,t+1} \tag{15}$$

$$V_{s,t} = \frac{\alpha n_b}{N_t}\theta_{sb}\mathbb{E}\pi + \frac{\alpha(n_{m,t} - n_t)}{N_t}\theta_{sm}R_t + \beta V_{s,t+1} \tag{16}$$

$$V_{0,t} = \frac{\alpha n_{s,t}}{N_t}\theta_{ms}R_t + \beta V_{0,t+1} \tag{17}$$

$$V_{1,t} = \rho + \frac{\alpha n_b}{N_t}\theta_{mb}\int_{R_t}^{\infty}(\pi - R_t)\,dF(\pi) + (1-\delta)\beta V_{1,t+1} + \delta\beta V_{0,t+1}. \tag{18}$$

Subtracting (17) from (18) and simplifying, we get a difference equation analogous to the differential equation (8),

$$R_{t-1} = (1-\delta)\,\beta\left\{\rho + R_t + \frac{\alpha n_b\theta_{mb}}{N_t}\int_{R_t}^{\infty}[1 - F(\pi)]\,d\pi - \frac{\alpha n_{s,t}\theta_{ms}}{N_t}R_t\right\}. \tag{19}$$

Similarly, we get a law of motion analogous to (9),

$$n_{t+1} = (1-\delta)\,n_t\left[1 - \frac{\alpha n_b}{N_t}\mathbb{E}\tau(\pi, R)\right] + (1-\delta)\,\frac{(n_{m,t} - n_t)\alpha n_{s,t}}{N_t}. \tag{20}$$

With no entry, one can check $dR_t/dR_{t-1} > 1$. Hence (19) has a unique equilibrium, which is the steady state R. Also, $dn_{t+1}/dn_t \in (0,1)$, so n_t converges to the steady state n. Now consider entry by S with a per-period cost, which reduces to exactly (10) in the benchmark model. Given initial n_0, equilibrium is a nonnegative, bounded path for (n_t, R_t) satisfying (19)-(20), written compactly as

$$\begin{bmatrix} R_{t-1} \\ n_{t+1} \end{bmatrix} = \begin{bmatrix} f(n_t, R_t) \\ g(n_t, R_t) \end{bmatrix}.$$

Now the n locus satisfying $n = g(n, R)$ and the R locus satisfying $R = f(n, R)$ both slope up in (n, R) space indicating the possibility of multiplicity.

Example 6: The same as Example 1 plus $\rho = 0.2$.

There are three steady states $(0.9007, 0.4213)$, $(1, 0.4421)$ and $(1.4826, 0.4777)$, similar to the continuous time specification. However, the discrete time dynamics are rather different. Let us focus on a two-cycle, oscillating between a liquid regime with low R and an illiquid regime with high R, denoted $\left(R^L, n^L\right)$ and (n^I, R^I). These solve

$$\left[\begin{array}{c} R^I \\ n^I \end{array}\right] = \left[\begin{array}{c} f(n^L, R^L) \\ g(n^L, R^L) \end{array}\right] \text{ and } \left[\begin{array}{c} R^L \\ n^L \end{array}\right] = \left[\begin{array}{c} f(n^I, R^I) \\ g(n^I, R^I) \end{array}\right]. \tag{21}$$

A solution is $\left(R^L, n^L\right) = (0.9800, 0.4511)$ and $\left(R^H, n^H\right) = (1.0065, 0.4297)$, shown in Fig. 1b. Fig. 6 shows the times series. In the liquid regime R is low, making M more likely to trade with B, and n is high because M and B traded less last period, while n_s is low because low R and high n discourage entry by S. The illiquid regime is just the opposite.[16]

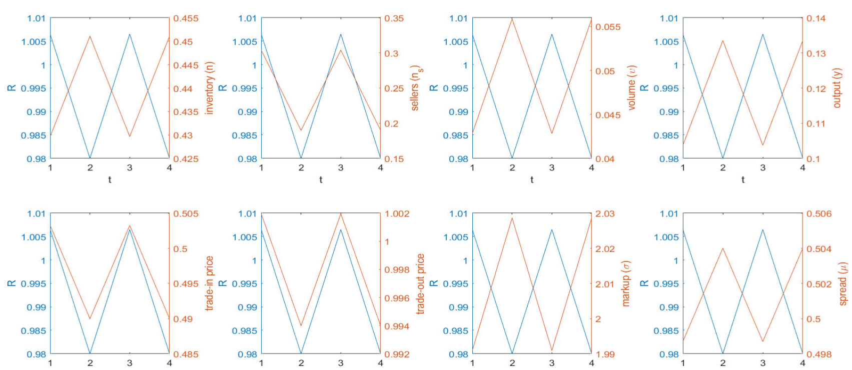

Fig. 6b: Example 6, Time Series.

Next, consider entry by M. Then (15)-(20) are the same, but n_s is fixed while

[16] Prices are also shown in Fig. 6b. The direct price is constant over time, depending only on fundamentals, but the wholesale and retail prices move with R. The spread can go either way, but here it moves against R. This is all broadly consistent with the data discussed in Comerton-Forde et al. (2010), and other stylized facts like inventories being more volatile than output. While this is, again, obviously not a calibration, the finding that it is qualitatively consistent with observations may lend further credence to the story.

$n_{m,t}$ is endogenous. Now (16) yields N_t in terms of R_t,

$$\kappa_m N_t = \alpha n_s \theta_{ms} R_t. \tag{22}$$

From (22), N_t depends only on R_t, while with entry by S it depends on R_t and n_t. Substituting (22) into (19), after some algebra we get $R_{t-1} = G(R_t)$, where

$$G(R) \equiv \beta(1-\delta)\left\{\rho + R + \frac{\kappa_m n_b \theta_{mb}}{n_s \theta_{ms} R}\int_R^\infty [1 - F(\pi)]\, d\pi - \kappa_m\right\}. \tag{23}$$

Now R_{t-1} depends only on R_t, while with entry by S it depends on R_t and n_t.

The univariate system $R_{t-1} = G(R_t)$ determines the path for R_t, from which we get N_t, n_t, etc. Steady state solves $R = G(R)$ as long as it implies $n_m, n \geq 0$, both of which hold iff $R \geq \underline{R} \equiv (n_s + n_b)\kappa_m/\alpha n_s \theta_{ms}$ (we also need $n \leq n_m$ but that never binds). A solution to $R = G(R) \geq \underline{R}$ is a steady state with M active. One can check $G(0) = \infty$, $G'(R) < 1$ and $G''(R) \geq 0$. Also, $\forall R > \bar{\pi}$ G is linear with slope $\beta(1-\delta)$. This is shown in Fig. 7a, from which it is clear that there exists a unique fixed point $\hat{R}$. We can have $\hat{R} > \bar{\pi}$, on the linear part of $G(R)$, or $\hat{R} < \bar{\pi}$, on the nonlinear part. If $G'(\hat{R}) < -1$ then standard methods imply there are cycles. There is a threshold ρ_1 such that $G'(\hat{R}) < -1$ iff $\rho < \rho_1$ (we do not know if $\rho_1 > 0$ or $\rho_1 < 0$ in general, but always found $\rho_1 < 0$ in examples). We now show that $G'(\hat{R}) < -1$ and $\hat{R} \leq \underline{R}$ are possible.

Example 7: $\pi \sim [0, 0.7]$, $\alpha = 1$, $\delta = 0.01$, $\beta = 0.99$, $n_b = n_s = 1$, $\theta_{mb} = 1$, $\theta_{ms} = 0.1$, $\kappa_m = 0.001$, and various ρ.

Fig. 7a depicts $G(R)$ in Example 7. As ρ decreases, the slope at steady state falls. One can check $G'(\hat{R}) < -1$ and $\hat{R} < \underline{R}$ when $\rho = -0.1$. Hence there is a 2-cycle and possibly cycles of higher order. The right panel of Fig. 7b plots the second and third iterates, $G^2(R)$ and $G^3(R)$. A fixed point of G^2 (G^3) other than a steady state is a 2-cycle (3-cycle). As shown, there exist a pair of 3-cycles. The existence of 3-cycles implies the existence of k-cycles $\forall k$ plus chaotic dynamics, by the Sarkovskii and Li-Yorke theorems.

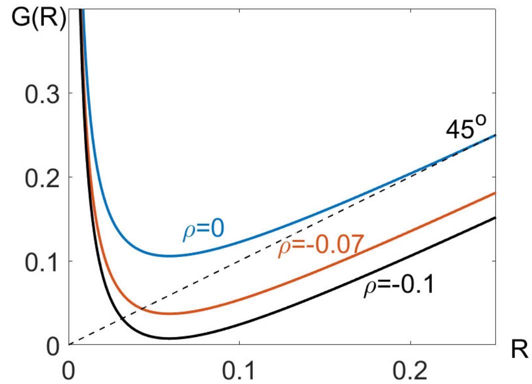

Fig.7a: Example 7, Different ρ.

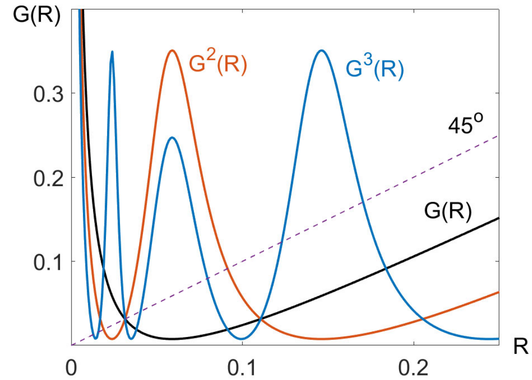

Fig.7b: Example 7, 2 and 3 Cycles, $\rho = -0.1$.

Therefore, this discrete-time model has many dynamic equilibria, and is actually easy to analyze, at least with entry by M, which implies a univariate system. Moreover, multiplicity and dynamics emerge with entry by M, counter to our intuition about having entry by one type and a different decision by another type. However, these equilibria vanish as the period length shrinks:

Proposition 3 *In the discrete-time model with entry by M, where h denotes the length of a period, there exists $\underline{h} > 0$ such that for all $h \in (0, \underline{h})$ no cycles exist.*

Oberfield and Trachter (2012), Rocheteau and Choi (2021) and Rocheteau and Wang (2023) in different models show cycles vanish as the discrete period length shrinks, motivating us to check which results are robust and which are not. While discrete time with entry by M is tractable and delivers interesting results, one might worry this is an artifact of the period length. Discrete time with entry by S is less tractable, but more robust: interesting dynamic equilibria still exist when the period length shrinks, as Section 3 shows – i.e., working directly with continuous time we established the existence of cyclic equilibria.

4.4 Inventories Without Middleman

Middleman are not necessary for our results; what actually matters is that there are both entry and inventory decisions. The middlemen framework is a very natural way to capture this, with entry decisions by S and inventory/trading decisions by M. Still, an environment can be designed with no type M, so S and B must trade directly, but now B has the option to consume x for payoff π or store it for return ρ, which can play the role of M's option to trade x or store it in the main model. One can interpret storage by B as inventory, or savings, as opposed to consumption. This inventory/savings option is only viable when $\rho > 0$, which is a reason one might prefer the specification with M, where interesting results can occur for $\rho \leq 0$ or $\rho > 0$. In any event, to be clear, the purpose of this extension is to show it is *possible* to get similar results without M, even if the model with M is better on some dimensions.

As usual B's payoff is match specific, $\pi \sim F(\pi)$, and B and S trade as long as $\pi > 0$. If B chooses not to consume x, it is inventoried for flow return ρ and depreciates at rate δ. Assume for simplicity that if B decides to store x the decision is irreversible – it is not possible to later consume it and go back on the market, so B is off the market until x depreciates. This restriction does not bind in, but could bind out of, steady state. While it is not especially natural, it lets us make the point relatively easily. Note that no such restriction is needed in the model with M, which may be another reason to prefer that version.

Normalize the measure of B to 1, and let S enter the market by paying κ_s. Let V_s, V_0 and V_1 be the value functions of S, B without inventory and B with inventory. When trading with S, if B inventories x the surplus is $\Delta = V_1 - V_0$, and if B consumes x the surplus is π. Then B consumes x if π exceeds the reservation value$R = \Delta$. If θ is B's bargaining power, then

$$
\begin{aligned}
rV_s &= \frac{\alpha(1-n)}{1+n_s}(1-\theta)\left[\int_0^R \Delta dF(\pi) + \int_R^{\bar{u}} \pi dF(\pi)\right] + \dot{V}_s \\
rV_0 &= \frac{\alpha n_s}{1+n_s}\theta\left[\int_0^R \Delta dF(\pi) + \int_R^{\bar{u}} \pi dF(\pi)\right] + \dot{V}_0 \\
rV_1 &= \rho - \delta\Delta + \dot{V}_1.
\end{aligned}
$$

This leads to

$$
\dot{\Delta} = (r+\delta)\Delta - \rho + \frac{\alpha n_s}{1+n_s}\theta\left\{\Delta + \int_\Delta^{\bar{u}} [1-F(\pi)]\, d\pi\right\},
$$

after integration by parts. The law of motion for inventories is

$$\dot{n} = -\delta n + \frac{\alpha (1-n) n_s}{1+n_s} F(\Delta)$$

and the entry condition by S implies

$$\frac{\alpha (1-n)}{1+n_s} (1-\theta) \left\{ \Delta + \int_{\Delta}^{\bar{u}} [1 - F(\pi)] \, d\pi \right\} = \kappa.$$

Consider first a degenerate distribution of π. That leads to

$$\begin{aligned} \dot{\Delta} &= (r+\delta)\Delta - \rho + \frac{\alpha n_s}{1+n_s}\theta [\gamma\pi + (1-\gamma)\Delta] \\ \dot{n} &= -\delta n + \frac{\alpha (1-n) n_s}{1+n_s} (1-\gamma) \\ \kappa &= \frac{\alpha (1-n)}{1+n_s} (1-\theta) [\gamma\pi + (1-\gamma)\Delta] \end{aligned}$$

where γ denotes B's probability of consuming x. In terms of steady state, there are three regimes with $n_s > 0$ so the market does not shut down, $\gamma = 1$, $\gamma = 0$ and $\gamma \in (0,1)$, plus a regime with $n_s = 0$. In the Appendix we construct the set of parameters that make γ a best response to itself, and check whether $n_s > 0$ or $n_s = 0$, as well as $n \in [0,1]$.

Example 8: $\alpha = 0.1$, $\delta = 0.01$, $n_b = 1$, $\theta = 0.73$, $\kappa_s = 0.01$ different π and ρ.

The results are shown in Fig. 8a, partitioning parameter space into 4 regions that support the different regimes. The pattern is general, and its properties are derived as Claim 1 in the Appendix, while the picture is drawn for the specification in Example 8. In the gray area $n_s = 0$, so the market shuts down. Otherwise, $n_s > 0$ and: in the blue region $\gamma = 1$ since π is high relative to ρ; in the brown region $\gamma = 0$ since π is low relative to ρ; in the green region there are three steady states, $\gamma = 0$, $\gamma = 1$ and $\gamma \in (0,1)$. Hence multiple steady states can exist, as in the main model, with M, with a similar intuition: if γ is high B is often without x, making it easier for S to trade, making n_s big and B more inclined to consume x; but if γ is low, and so on.[17]

[17] It is tempting so suggest that this has a Keynesian flavor, with γ denoting the marginal propensity to consume. It does have the feature that more consumption (higher aggregate demand) stimulates more production (higher aggregate supply) as a self-fulfilling prophecy.

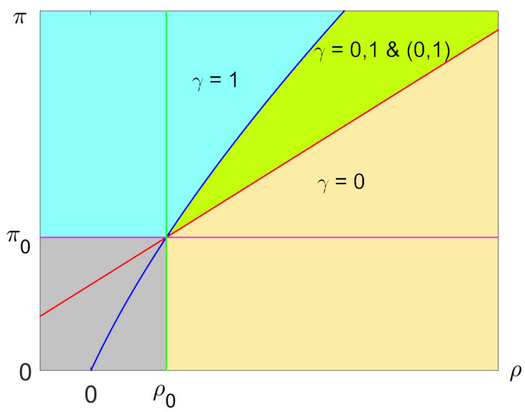

Fig. 8a: Regions/regimes without M.

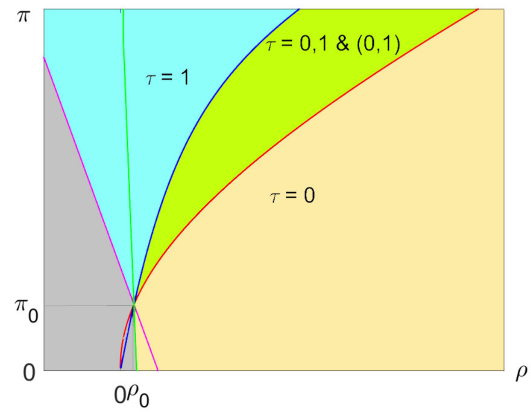

Fig. 8b: Regions/regimes with M.

If Fig. 8a is useful, Fig. 8b provides a similar picture for the model with M, drawn for Example 1 in Section 3. However, the interpretation is: here S and B always trade; S and M do not trade the grey area and trade in other regions; and those other regions differ in the probability τ that M trades with B. We did not draw this graph earlier because deriving the regions with three types is more cumbersome – Fig. 8b is done numerically – and because these pictures are only relevant for degenerate π, while the preferred specification has disperse π. Still, the figures are remarkably similar, and Fig. 8b nicely tightens a loose end in Section 3, where it was simply assumed that ρ is above some lower bound to guarantee M and S trade; now the boundary of grey area tells us just how low ρ can go before S and stop M trading. More generally, these pictures indicate that steady state exists for all parameters, although for some parameters it is possible that certain agents stop trading.[18]

Example 9: π is as in (14) with $a = 0.3$, $b = 0.31$, $c = 0.48$, $d = 0.5$, $y_1 = 0.01$, $y_2 = 0.49$, $y_3 = 0.5$, $\alpha = 0.46$, $\delta = 0.046$, $n_b = 1$, $\theta = 0.75$, $\kappa_s = 0.01$ and $\rho = 0.1$.

[18] In fact two things can go wrong when trying to construct an equilibrium where S and M trade: we can violate $\Delta \geq 0$ or $n \geq 0$. In Fig. 8b, what binds is $n \geq 0$ when we hit the grey area as we lower ρ for small π, since we hit it before we reach $\rho = 0$ and $\rho > 0$ means $\Delta > 0$. For higher π what binds is $\Delta \geq 0$ when we hit the grey area as we lower ρ.

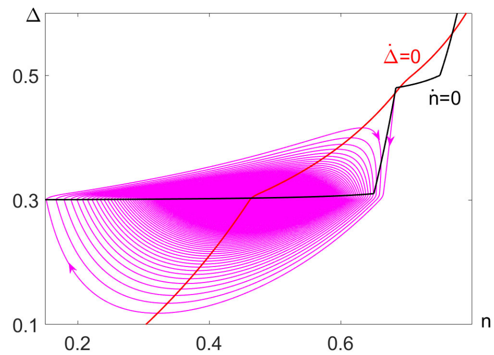

Fig. 9a: Example 9, $r = 0.015$.

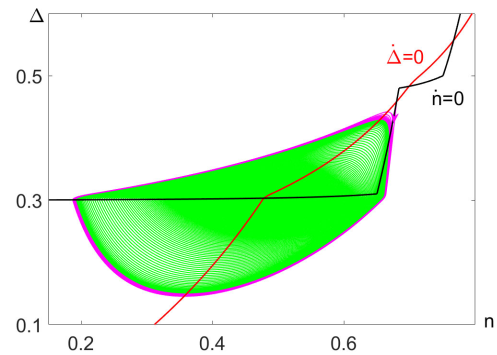

Fig. 9b: Example 9, $r = 0.020$.

Going beyond steady state, the model without M also has cyclic equilibria. As r varies in Example 9 there is a Hopf bifurcation at $r^* = 0.0174$. Fig. 9a shows $r = 0.015$, where there are four steady states, and the lowest is a sink, and the unstable manifold of the next lowest steady state converges to the lowest one. As r increases past r^* the sink becomes a source, with a stable limit cycle around it. Fig. 9b shows the case with $r = 0.020$, with the green curve showing a trajectory spiraling away from the source and converging to a cycle. The unstable manifold also converges to a cycle. Since cycles emerge when r increases, $r^*$is supercritical.

Finally, we check on the general intuition offered above, that multiplicity and cycles emerge when agents on one side make a inventory or savings decision while those on the other side make an entry decision. The Appendix considers a version of this model, without M, where B makes both an entry decision and inventory decision and proves as Claim 2 that equilibrium is unique. The Appendix also considers a version, without M, where S makes both an entry decision and inventory decision and similarly proves as Claim 3 that equilibrium is unique. This is all consistent with our general intuition.

5 Conclusion

This paper studied dynamic models of inventories, focusing on intermediated trade, and allowing heterogeneous buyer valuations. We showed there are multiple steady states, and endogenous cycles, where entry, trading strategies, liquidity, prices and other endogenous variables fluctuate. These fluctuations are driven by strategic considerations, not increasing returns, as in some other models, or the self-referential nature of acceptability, as in monetary economics. They are possible for inventories with a positive return or a storage cost – i.e., for asset or goods markets. For asset markets, this is consistent with the venerable view that financial intermediaries are prone to instability or volatility, while for goods markets, it is consistent with the fact that retail trade differs dramatically across economies, as discussed in the Introduction. We analyzed discrete- and continuous-time specifications. In some cases (entry by middlemen) discrete time was tractable and gave interesting results, but they are not robust to period length; in other cases (entry by sellers), cyclic equilibria are possible in discrete and continuous time.

As mentioned above, the relevance of the findings is this: while it may be hard to account for data based purely on self-fulfilling prophecies, when simple and natural models display such outcomes, it may make one more inclined to think that actual economies can, too. This is consistent with Diamond's (1982b) view, but to get the desired results he needed increasing returns, which he called an externality. As he said, "this externality involves positive feedback: increased production for inventory makes trade easier; easier trade makes production for inventory more profitable and therefore justifies its increase. This positive feedback... implies the possibility of multiple equilibria." His models do not capture inventory behavior the way we do – there are no middlemen – but, interpreted broadly, the spirit is similar.

One might say that when strategic considerations that arise naturally with intermediation are introduced, we can dispense with increasing returns and still get multiplicity and complicated dynamics. To put this in context, Diamond (1984) got money into his framework with a CIA (cash-in-advance) restriction, and again increasing returns led to multiplicity. Subsequent developments showed that modeling the microfoundations in more detail means the CIA constraint is not needed to get valued fiat currency, and that further implies increasing returns are not needed to get natural multiplicities and dynamics in monetary economics, as emphasized by, e.g., Kiyotaki and Wright (1993) or Johri (1999). What we think is a general conclusion is that markets with frictions are prone to volatility or instability, in the sense that there can be multiple steady states and cyclic dynamic equilibria arising as self-fulfilling prophecies. This is well know in monetary models. Our results show that real models with inventories have similar properties.

In other words, once the exchange process is modeled in more detail, and exchange is not as simple as it is in Diamond's early work, where everyone trade with everybody, there is a role for institutions that facilitate this process. Two such institutions are money and middlemen. Once they are modeled explicitly, one can dispense with mechanical assumptions like increasing returns and get similarly interesting results.

Appendix

Proof of Proposition 1: First notice that, with $\pi = \bar{\pi}$, (8) and (9) reduce to

$$\begin{aligned}(r+\delta+\alpha\theta_{ms})\Delta - \rho - \frac{\alpha n_b \theta_{mb} \tau (\bar{\pi}-\Delta) + \alpha (n_b+n_m)\theta_{ms}\Delta}{N} &= 0 \quad (24)\\ \delta n + \frac{\alpha n n_b \tau}{N} - \alpha(n_m - n)\left(1 - \frac{n_b+n_m}{N}\right) &= 0 \quad (25)\end{aligned}$$

where

$$N = \frac{\alpha n_b \theta_{sb}\bar{\pi} + \alpha (n_m - n)\theta_{sm}\Delta}{\kappa_s}.$$

In the region where $\Delta > \bar{\pi}$, where $\tau = 0$, combine (24) and (25) to eliminate N,

$$\left(r+\delta+\frac{\theta_{ms}\delta n}{n_m - n}\right)\Delta = \rho. \tag{26}$$

This implies

$$\frac{\partial \Delta}{\partial n} = -\frac{\theta_{ms}\delta n_m \Delta}{(n_m-n)^2 (r+\delta) + (n_m - n)\theta_{ms}\delta n} < 0.$$

Thus we transform system (24)-(25) to (25)-(26). As (26) is downward sloping and (25) upward sloping, there is at most one steady state with $\Delta > \bar{\pi}$. Also, from (26), steady state exists in this region only if $\rho > 0$.

In the region where $\Delta < \bar{\pi}$, where $\tau = 1$, combine (24) and (25) to get

$$(r+\delta+\alpha\theta_{ms})\Delta = \rho + \frac{n_b\theta_{mb}(\bar{\pi}-\Delta) + (n_b+n_m)\theta_{ms}\Delta}{n_m(n_b+n_m-n)}\left[\alpha(n_m-n) - \delta n\right].$$

This implies

$$\frac{\partial \Delta}{\partial n} = -\frac{n_b\theta_{mb}(\bar{\pi}-\Delta)+(n_b+n_m)\theta_{ms}\Delta}{r+\delta+\frac{\theta_{ms}n[\alpha n_b+\delta(n_m+n_b)]+\theta_{mb}n_b[\alpha(n_m-n)-\delta n]}{n_m(n_b+n_m-n)}}\frac{\delta(n_b+n_m)+\alpha n_b}{n_m(n_b+n_m-n)^2} < 0. \tag{27}$$

Again, since (27) is downward and (25) upward sloping, there is at most one steady state with $\Delta < \pi$. Similarly, when $\Delta = \bar{\pi}$ and $\tau \in (0,1)$, the n locus is flat and Δ locus upward sloping. Hence, there again is at most one steady state.

For existence, it is easily verified that Δ and n loci are upward sloping and the Δ locus is flatter than the n locus in regions where $\Delta \neq \bar{\pi}$. Also, the Δ locus shifts up when ρ increases. For $\rho = -\theta_{mb}\kappa_s/\theta_{sb}$, the Δ locus goes through the origin. For

$\rho > -\theta_{mb}\kappa_s/\theta_{sb}$, the Δ locus has a positive intercept. At $n = n_m$, Δ is positive and finite on the Δ locus. The n locus goes through $(\underline{n}, 0)$, where

$$\underline{n} \equiv \alpha n_m \left(1 - \frac{n_b + n_m}{\alpha n_b \theta_{sb} \bar{\pi}} \kappa\right) / \left[\delta + \frac{\kappa}{\theta_{sb}\bar{\pi}} + \alpha \left(1 - \frac{n_b + n_m}{\alpha n_b \theta_{sb}\bar{\pi}} \kappa\right)\right],$$

strictly first increases, then becomes flat and goes to ∞ as n goes to $\alpha n_m / (\alpha + \delta) < n_m$. Hence the loci have at least one intersection. In particular, if there is a steady state at $\Delta = \bar{\pi}$, there are two more steady states, one with $\Delta < \bar{\pi}$ and one with $\Delta > \bar{\pi}$. As ρ shifts the Δ locus, there exist $\tilde{\rho}, \hat{\rho} \geq 0$ with the stated properties. ■

Proof of Proposition 2. Suppose M buys x from S and does not sell it to B. If $\rho < 0$ this "buy and hold" strategy has a negative payoff which is dominated by not buying x. ■

Proof of Proposition 3. Let the length of a period be h. Then β, α, κ_m, ρ and δ are functions of h. As usual, let:

$$r = \lim_{h \to 0} \frac{\beta(h)^{-1} - 1}{h}, \ \alpha = \lim_{h \to 0} \frac{\alpha(h)}{h}, \ \kappa_m = \lim_{h \to 0} \frac{\kappa_m(h)}{h}, \ \rho = \lim_{h \to 0} \frac{\rho(h)}{h}, \ \delta = \lim_{h \to 0} \frac{\delta(h)}{h}$$

The equilibrium condition can be rewritten $R_{t-h} = G(R_t, h)$, where

$$G(R; h) \equiv \beta(h) [1 - \delta(h)] \left\{\rho(h) + R + \frac{\kappa_m(h) n_b \theta_{mb}}{n_s \theta_{ms} R} \int_R^\infty [1 - F(\pi)] d\pi - \kappa_m(h)\right\}. \tag{28}$$

As $h \to 0$, this converges to the continuous-time model.

First we show steady state is unique. From (28) we get

$$\frac{\partial G}{\partial R} = \beta(h) [1 - \delta(h)] \left\{1 + \frac{\kappa_m(h) n_b \theta_{mb}}{n_s \theta_{ms} R} \left[-1 + F(R) - \frac{1}{R} \int_R^\infty [1 - F(\pi)] d\pi\right]\right\}. \tag{29}$$

As the term in square brackets is negative, $\partial G / \partial R < \beta(h) [1 - \delta(h)] < 1 \ \forall R$, so G crosses the 45^o line at most once: if it exists steady state is unique. With $\hat{R}(h)$ denoting steady state as a function of period length, it solves $R = G(R, h)$, which we rearrange as

$$\begin{aligned} & R^2 [1 - \beta(h) - \beta(h) \delta(h)] \\ = \ & \beta(h) [1 - \delta(h)] \left\{[\rho(h) - \kappa_m(h)] R + \frac{\kappa_m(h) n_b \theta_{mb}}{n_s \theta_{ms}} \int_R^\infty [1 - F(\pi)] d\pi\right\}. \end{aligned}$$

For any $h > 0$, the LHS is 0 at $R = 0$ and the RHS is strictly positive at $R(h) = 0$. Hence $\hat{R}(h) > 0 \; \forall h > 0$. Dividing by h we get

$$\begin{aligned} & R^2 \frac{1 - \beta(h) - \beta(h)\delta(h)}{h} \\ = \; & \beta(h)[1 - \delta(h)] \left\{ \frac{\rho(h) - \kappa_m(h)}{h} R + \frac{\kappa_m(h)}{h} \frac{n_b \theta_{mb}}{n_s \theta_{ms}} \int_R^\infty [1 - F(\pi)] \, d\pi \right\}. \end{aligned}$$

As $h \to 0$, the LHS approaches $R^2(r + \delta)$ and the RHS approaches $(\rho - \kappa_m) R + \frac{\kappa_m n_b \theta_{mb}}{n_s \theta_{ms}} \int_R^\infty [1 - F(\pi)] \, d\pi$. At $R = 0$, the LHS is 0 and the RHS is strictly positive. Hence, $\hat{R}(0) \neq 0$. Finally, evaluate (29) at $h = 0$ and $\hat{R}(0)$ to get $\lim_{h \to 0} \partial G / \partial \hat{R}_0 = 1$. By the continuity of G, there exists a cutoff $\underline{h} > 0$ such that $\partial G / \partial \hat{R}(h) > -1$ for $h > \underline{h}$, implying a 2-cycle, and $\partial G / \partial \hat{R}(h) \leq -1$ otherwise. As is standard, if a 2-cycle does not exist, no cycles of any order exists. ■

Proof of Claim 1: As discussed in the text there are 4 cases. We consider each in turn.

Case 1: $\gamma = 1$ and $n_s > 0$. For $\gamma = 1$ we need $\pi > \Delta$, which easily reduces to

$$\pi \geq f_1(\rho) \equiv \frac{\rho + \kappa\theta / (1 - \theta)}{r + \delta + \alpha\theta} \tag{30}$$

Given $\gamma = 1$, $n = 0$. Finally, $n_s > 0$ reduces

$$\pi \geq \frac{\kappa}{\alpha(1 - \theta)} \equiv \pi_0 \tag{31}$$

Case 2: $\gamma = 0$ and $n_s > 0$. Now $\gamma = 0$ requires $\Delta \geq \pi$, which reduces to

$$\pi < \max\{f_2(\rho), f_3(\rho)\} \tag{32}$$

where $f_2(\rho) \equiv \rho / (r + \delta + \alpha\theta)$ and

$$\begin{aligned} f_3(\rho) \equiv \; & \frac{\delta(1 - \theta)(\rho - \kappa) - \kappa r}{2(1 - \theta)\delta(r + \delta + \alpha\theta)} \\ & + \frac{\left\{[\delta(1 - \theta)(\rho - \kappa) - \kappa r]^2 + 4\delta(1 - \theta)(r + \delta + \alpha\theta)\kappa\rho\right\}^{0.5}}{2(1 - \theta)\delta(r + \delta + \alpha\theta)}. \end{aligned}$$

Notice f_3 is increasing, concave, $f_3(0) = 0$ and $f_3(\rho_0) = \pi_0$ where π_0 is given in(31). Then $n \geq 0$ reduces to

$$\rho > \frac{\kappa(r + \delta)}{\alpha(1 - \theta)} \equiv \rho_0. \tag{33}$$

We also need $n \leq 1$ and $n_s > 0$ but these are redundant given the other conditions.

Case 3: $\gamma \in (0,1)$ and $n_s > 0$. For $\gamma \in (0,1)$ we solve for $\Delta = \pi$ for n and check that $0 \leq n \leq 1$ holds iff $\pi \geq f_1(\rho)$ and $0 < \gamma < 1$ iff $\pi < f_3(\rho)$, which also guarantees $n_s > 0$. Notice that when two pure-strategy steady states exist, as usual, the mixed-strategy steady state does, too.

Case 4: $n_s = 0$. We need $rV_s \leq \kappa$ which reduces to a simple parameter condition. Here obviously $n = 0$ so we only need to check the best response condition for γ, even if it is only relevant off the equilibrium path, since $n_s = 0$ means B never actually gets to decide to consume or store x in equilibrium. For any $\gamma \in [0,1]$, $\Delta = \rho/(r+\delta)$ as B's trading probability is 0. For $\gamma = 1$, $\Delta < \pi$ iff $\pi > \rho/(r+\delta)$ and $rV_s \leq \kappa$ iff $\pi \leq \pi_0$. For $\gamma = 0$, $\Delta > \pi$ iff $\pi < \rho/(r+\delta)$ and $rV_s \leq \kappa$ iff $\rho \leq \rho_0$. For $\gamma \in (0,1)$, $\pi = \rho/(r+\delta)$ and $rV_s \leq \kappa$ iff $\pi \leq \pi_0$. Altogether, $n_s = 0$ holds iff $\pi \leq \pi_0$ and $\rho \leq \rho_0$. ■

Proof of Claim 2: Consider the model without M where B decides whether to enter and whether to consume or store x. We have

$$\begin{aligned}
rV_s &= \frac{\alpha(n_b - n)}{1+n_b}(1-\theta)\left[\int_0^R \Delta dF(\pi) + \int_R^{\bar{u}} \pi dF(\pi)\right] + \dot{V}_s \\
rV_0 &= \frac{\alpha}{1+n_b}\theta\left[\int_0^R \Delta dF(\pi) + \int_R^{\bar{u}} \pi dF(\pi)\right] + \dot{V}_0 \\
rV_1 &= \rho - \delta\Delta + \dot{V}_1
\end{aligned}$$

Following the usual procedure, we get

$$\begin{aligned}
\dot{\Delta} &= (r+\delta)\Delta - \rho + \frac{\alpha}{1+n_b}\theta\left\{\Delta + \int_\Delta^{\bar{u}} [1 - F(\pi)]\, d\pi\right\} \\
\dot{n} &= -\delta n + \frac{\alpha(n_b - n)}{1+n_b}F(\Delta) \\
\kappa_b &= \frac{\alpha}{1+n_b}\theta\left[\int_0^R \Delta dF(\pi) + \int_R^{\bar{u}} \pi dF(\pi)\right]
\end{aligned}$$

Again, using the third equation to eliminate n_b from the others we have a bivariate system

In particular, the entry condition implies

$$\dot{\Delta} = (r+\delta)\Delta - \rho + \kappa_b,$$

which has a unique bounded solution, the steady state, $\Delta = (\rho - \kappa) / (r + \delta)$. Then

$$n_b = \frac{\alpha}{\kappa_b}\theta \left[\int_0^R \Delta dF(\pi) + \int_R^{\bar{u}} \pi dF(\pi) \right] - 1$$

is also a constant, in and out of steady state, while inventories converge to steady state following the $\dot{n}$ equation. Hence there is a unique equilibrium. ■

Proof of Claim 3: Now suppose S decide whether to produce for themselves at cost c and enjoy ρ, or enter the market and produce when they meet B at the same cost c, or not produce. Let V_0 and V_1 be the value functions of S without and with inventory, and V_s the value of S in the market. Let N_s be the total measure of sellers and n_s those in the market. For simplicity, suppose π is degenerate. Then

$$\begin{aligned} V_0 &= \max\{V_1 - c, V_s, 0\} \\ rV_1 &= \rho - \delta(V_1 - V_0) + \dot{V}_1 \\ rV_s &= \frac{\alpha n_b}{n_s + n_b}\theta_s[\pi - c - (V_s - V_0)] + \dot{V}_s \end{aligned}$$

Now we can proceed as in Claim 1 and consider four regimes, although now we do not restrict attention to steady state.

Case 1: All sellers enter. That requires $V_0 = V_s > V_1 - c$ and $n_s = N_s$. There is a unique solution

$$V_s = \frac{\alpha n_b \theta_s}{r(N_s + n_b)}(u - c) \text{ and } V_1 = \frac{\rho + \delta V_s}{r + \delta},$$

and this is an equilibrium for parameters satisfying $u > c$ and

$$\rho \leq f_1(u) \equiv \frac{\alpha n_b \theta_s}{n_b + N_s} u + \left(r + \delta - \frac{\alpha n_b \theta_s}{n_b + N_s} \right) c.$$

Case 2: All sellers produce for themselves. That requires $V_0 = V_1 - c \geq V_s$ and $n_s = 0$. There is again a unique solution

$$V_1 = \frac{\rho - \delta c}{r} \text{ and } V_s = \frac{\alpha\theta_b}{r + \alpha\theta_s}(u - 2c + V_1)$$

and this is an equilibrium for parameters satisfying $\rho \geq f_2(u) \equiv \alpha\theta_s u + (r + \delta - \alpha\theta_s) c$ and $\rho > (r + \delta) c$.

Case 3: Some sellers produce and hold inventory, and others enter the market. That requires $V_0 = V_s = V_1 - c$ and $n_s \in (0, N_s)$. The unique solution is

$$V_1 = \frac{\rho - \delta c}{r}, \; V_s = \frac{\rho - (r + \delta)\, c}{r}, \text{ and } n_s = \frac{\alpha n_b \theta_s (u - c)}{\rho - (r + \delta)\, c} - n_b$$

and this is an equilibrium for parameters satisfying $f_1(u) < \rho < f_2(u)$.

Case 4: Sellers do not produce. That requires $V_0 = 0$, $V_1 - c < 0$, $V_s < 0$ and $n_s = 0$. The unique solution is

$$V_1 = \frac{\rho}{r + \delta} \text{ and } V_s = \frac{\alpha \theta_s (u - c)}{r + \alpha \theta_s}$$

and this is an equilibrium for parameters satisfying $\rho < (r + \delta)\, c$ and $u < c$. This completes all the cases, and implies uniqueness. ■

References

J. Albrecht, F. Postel-Vinay and S. Vroman (2013) "An Equilibrium Search Model of Synchronized Sales," *IER* 54, 473-93.

C. Azariadis (1993) *Intertemporal Macroeconomics*.

J. Benhabib and K. Nishimura (1979) "The Hopf Bifurcation and the Existence and Stability of Closed Orbits in Multisector Models of Optimal Economic Growth," *JET* 21, 421-44.

J. Benhabib, S. Schmitt-Grohe and M. Uribe (2001) "The Perils of Taylor Rules," *JET* 96 (2001), 40-69.

A. Blinder (1990) *Inventory Theory and Consumer Behavior*.

M. Boldrin N. Kiyotaki and R. Wright (1993) "A Dynamic Equilibrium Model of Search, Production, and Exchange," *JEDC* 17, 723-58.

M. Brennan, T. Chordia and A. Subrahmanyam (1998) "Alternative Factor Specifications, Security Characteristics, and the Cross-Section of Expected Stock Returns," *JFE* 49, 345-73.

K. Burdett and M. Coles (1997) "Marriage and Class," *QJE* 112, 141-168.

_____ (1998) "Separation Cycles," *JEDC* 22, 1069-90.

K. Burdett and R. Wright (1998) "Two-Sided Search with Nontransferable Utility," *RED* 1, 220-45.

R. Cavalcanti and N. Wallace (1999) "Inside and Outside Money as Alternative Media of Exchange," *JMCB* 31, 443-57.

M. Coles and R. Wright (1998) "A Dynamic Equilibrium Model of Search, Bargaining, and Money," *JET* 32-54.

C. Comerton-Forde, T. Hendershott, C. Jones, P. Moulton and M. Seasholes (2010) "Time Variation in Liquidity: The Role of Market-Maker Inventories and Revenues," *JF* 65, 295-331.

P. Diamond (1982*a*) "Aggregate Demand Management in Search Equilibrium," *JPE* 90, 881-94.

_____ (1982*b*) *A Search Equilibrium Approach to the Micro Foundations of Macroeconomics*.

_____ (1984) "Money in Search Equilibrium," *Econometrica* 52, 1-20.

P. Diamond and D. Fudenberg (1989) "Rational Expectations Business Cycles in Search Equilibrium," *JPE* 97, 606-19.

D. Duffie, N. Gârleanu and L. Pederson (2005) "Over-the-Counter Markets," *Econo-*

metrica 73, 1815-47.

M. Farboodi, G. Jarosch, G. Menzio and U. Wiriadinata (2019) "Intermediation as Rent Extraction," *NBER* WP 24171.

M. Farboodi, G. Jarosch and R. Shimer (2023) "The Emergence of Market Structure," *RES* 90, 261-92.

C. Fershtman and A. Fishman (1992) "Price Cycles and Booms: Dynamic Search Equilibrium," *AER* 82, 1221-33.

G. Gong (2023) "Middlemen in Search Equilibrium with Intensive and Extensive Margins," mimeo.

C. Gu, G. Menzio, R. Wright and Y. Zhu (2022) "Market Freezes," mimeo.

C. Gu, C. Monnet, E. Nosal and R. Wright (2023) "Diamond-Dybvig and Beyond: On the Instability of Banking," *EER*, in press.

J. Guckenheimer and P. Holmes (1983) *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields.*

P. Howitt and R. McAfee (1988) "Stability of Equilibria with Externalities," *QJE* 103, 261-77.

_____ (1992) "Animal Spirits," *AER* 82, 493-507.

J. Hugonnier, B. Lester and P. Weill (2020) "Frictional Intermediation in Over-the-Counter Markets," *RES* 87, 1432-69.

A. Johri (1999) "Search, Money, and Prices." *IER* 40, 439-54.

A. Khan and J. Thomas (2007) "Inventories and the Business Cycle: An Equilibrium Analysis of (S, s) Policies," *AER* 97, 1165-88.

G. Kaplan and G. Menzio (2016) "Shopping Externalities and Self-Fulfilling Unemployment Fluctuations," *JPE* 124, 771-825.

T. Kehoe N. Kiyotaki and R. Wright (1993) "More on Money as a Medium of Exchange," *ET* 3, 297-314.

N. Kiyotaki and R. Wright (1989) "On Money as a Medium of Exchange," *JPE* 97, 927-54.

_____ (1993) "A Search-Theoretic Approach to Monetary Economics," *AER* 83, 63-77.

Y. Kuznetsov (2004) *Elements of Applied Bifurcation Theory*, 2nd ed.

R. Lagos and G. Rocheteau (2009) "Liquidity in Asset Markets with Search Frictions," *Econometrica* 77, 403-26.

R. Lagos, G. Rocheteau and R. Wright (2017) "Liquidity: A New Monetarist Perspective," *JEL* 77, 403-26.

D. Lagakos (2016) "Explaining Cross-Country Productivity Differences in Retail Trade," *JPE.*

J. Martel, E. Van Wesep and B. Waters (2023) "Constrained Asset Prices," mimeo.

A. Masters (2007) "Middlemen in Search Equilibrium," *IER* 48, 343-62.

_____ (2008) "Unpleasant Middlemen," *JEBO* 68, 73-86.

D. Mortensen (1999) "Equilibrium Unemployment Dynamics," *IER* 40, 889-914.

E. Nosal, Y. Wong and R. Wright (2015) "More on Middlemen: Equilibrium Entry and Efficiency in Intermediated Markets," *JMCB* 47, 7-37.

_____ (2019) "Intermediation in Markets for Goods and Markets for Assets," *JET* 183, 876-906.

E. Oberfield and N. Trachter (2012) "Commodity Money with Frequent Search," *JET* 147, 2332-56.

B. Pierre (2019) "Let's Talk About Inventory Cycles," https://www.spscommerce.com/blog/talk-inventory-cycle-spsa/.

C. Pissarides (2000) *Equilibrium Unemployment Theory.*

J. Renero (1998) "Unstable and Stable Steady-States in the Kiyotaki–Wright Model," *ET* 11, 275-94.

G. Rocheteau and M. Choi (2021) "New Monetarism in Continuous Time: Methods and Applications," *EJ* 131, 658-96.

G. Rocheteau and E. Nosal (2017) *Money, Payments, and Liquidity.*

G. Rocheteau and L. Wang (2023) "Endogenous Liquidity and Volatility," *JET*

G. Rocheteau and R. Wright (2005) "Money in Search Equilibrium, in Competitive Equilibrium, and in Competitive Search Equilibrium," *Econometrica* 73, 175-202.

_____ (2013) "Liquidity and Asset-Market Dynamics," *JME* 60, 275-94.

A. Rubinstein and A. Wolinsky (1987) "Middlemen," *QJE* 102, 581-94.

F. Sniekers (2018) "Persistence and Volatility of Beveridge Cycles," *IER* 59, 665-98.

D. Vanyos and T. Wang (2007) "Search and Endogenous Concentration of Liquidity in Asset Markets," *JET* 136, 66-104.

P. Weill (2008) "Liquidity Premia in Dynamic Bargaining Markets," *JET* 140, 66-96.

R. Wright and Y. Wong (2014) "Buyers, Sellers and Middlemen: Variations on Search-Theoretic Themes," *IER* 55, 375-98.